Lithuanian Mathematical Journal, Vol. 52, No. 3, July, 2012, pp. 259–281

ESTIMATING ASYMPTOTIC DEPENDENCE FUNCTIONALS IN MULTIVARIATE REGULARLY VARYING MODELS

Georg Mainik

RiskLab, Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland
(e-mail: georg.mainik@math.ethz.ch)

Received September 14, 2011; revised May 24, 2012

Abstract. This paper deals with semiparametric estimation of the asymptotic portfolio risk factor γ_ξ introduced in [G. Mainik and L. Rüschendorf, On optimal portfolio diversification with respect to extreme risks, *Finance Stoch.*, 14:593–623, 2010] for multivariate regularly varying random vectors in $\mathbb{R}_+^d$. The functional γ_ξ depends on the spectral measure Ψ, the tail index α, and the vector ξ of portfolio weights. The representation of γ_ξ is extended to characterize the portfolio loss asymptotics for random vectors in $\mathbb{R}^d$. The earlier results on uniform strong consistency and uniform asymptotic normality of the estimates of γ_ξ are extended to the general setting, and the regularity assumptions are significantly weakened. Uniform consistency and asymptotic normality are also proved for the estimators of the functional $\gamma_\xi^{1/\alpha}$ that characterizes the asymptotic behavior of the portfolio loss quantiles. The techniques developed here can also be applied to other dependence functionals.

MSC: primary 60F17; secondary 60F05, 60F15, 60H12, 60G70

Keywords: tail dependence, multivariate regular variation, portfolio risk, functional CLT, functional SLLN

1 INTRODUCTION

Modeling dependence in multivariate random vectors and the estimation of various dependence characteristics is of importance in many applications of probability theory and statistics. In the areas of finance and insurance, any portfolio modeling effort includes dependence modeling and aggregation of random variables with nontrivial dependence structures. Beyond understanding portfolio behavior under benign market conditions, the sensitivity of the portfolio to extremal events deserves special attention. The aggregation of risk is also related to risk diversification, i.e., assessing the risk of a weighted sum $\sum_{i=1}^d \xi^{(i)} X^{(i)}$ for a random vector $X = (X^{(1)}, \dots, X^{(d)})$ and portfolio weights $\xi^{(1)}, \dots, \xi^{(d)}$.

A useful mathematical framework for modeling the extremal behavior of random vectors is multivariate extreme value theory. Starting with the characterization of joint distributions for componentwise maxima in [10], multivariate extreme value theory found many applications in insurance and finance (cf. [28, 30]). The literature on modeling and estimation of extremal dependence is vast, including several alternative approaches, such as tail dependence functions, spectral measures and alike, and extreme value copulas. See [9, 19, 20, 33], and references therein.

This paper presents a semiparametric approach to the estimation of tail dependence functionals for multivariate regularly varying models. Particular consequences of this assumption are that the loss components are heavy-tailed and that the extremal events for different components are on the same scale. This allows focusing on the cases with nontrivial contribution of dependence to the extremal behavior. The overall excess severity is characterized by the tail index $\alpha := \sup\{\beta \geqslant 0\colon \mathbf{E}\|X\|^{\beta} < \infty\}$, which is the number that separates the finite absolute moments from the infinite ones.

In nondegenerate cases, all components $X^{(i)}$ have the same tail index α. This computational advantage of multivariate regularly varying models is also their most critical issue. It is well known that in case of different tail indices the asymptotic distribution of the portfolio loss is dominated by the components with the heaviest tail, i.e., the one or the few $X^{(i)}$ with the smallest α. One should always bear this potential issue in mind when assuming multivariate regular variation. However, if the component tail indices are close to each other, the excess behavior over finite thresholds may be better described by a model with equally heavy component tails. Hence, from the practical point of view, multivariate regular variation can be understood as a workable approximation for the case of statistically indistinguishable component tail indices. Moreover, many popular models, such as heavy-tailed elliptical or multivariate α-stable distributions with $\alpha \in (0, 2)$, are multivariate regularly varying (cf. [2, 23]). Thus results obtained in this modeling framework contribute to the general understanding of asymptotic dependence.

The dependence structure in multivariate regularly varying models is characterized by the so-called spectral measure Ψ, which is the asymptotic probability distribution of excess directions on a unit sphere. In this modeling framework, the tail dependence coefficient and many other dependence characteristics can be represented in terms of integrals with respect to the spectral measure Ψ.

The estimation of dependence characteristics in the framework of extreme value theory or, more specifically, of multivariate regular variation, is by now a vital research field. The estimation of Ψ in the bivariate case was studied in [14]. A purely nonparametric approach to the estimation of dependence structures in the more general framework of multivariate extreme value theory was proposed in [15]. This setting was recently reconsidered for maximum likelihood estimation in [16]. A parametric approach to the estimation of the spectral measure of heavy-tailed elliptical distributions has been considered in [25]. Nonparametric estimation of tail dependence beyond the framework of extreme value theory has been studied in [34].

The contribution of the present paper to the estimation of extremal dependence is primarily focused on the asymptotic portfolio risk factor $\gamma_\xi = \gamma_\xi(\Psi, \alpha)$, which characterizes the influence of the tail dependence structure and the portfolio weights $\xi^{(1)}, \ldots, \xi^{(d)}$ on the extremal behavior of the portfolio risk [26, 27]. For related results we also refer to [3, 17]. The functional γ_ξ was originally defined for random vectors in $\mathbb{R}^d_+$, which is appropriate for applications in insurance. The initial approach is now extended to random vectors in $\mathbb{R}^d$, which allows accounting for losses and gains simultaneously. This is particularly useful in financial applications or in any other problem area that involves modeling the compensation between losses and gains.

The estimator $\widehat{\gamma}_\xi$ considered here is semiparametric, combining an appropriate estimator of the tail index $\alpha \in (0, \infty)$ with an empirical estimator of the spectral measure Ψ. The main results are the strong consistency and the asymptotic normality of $\widehat{\gamma}_\xi$ uniformly in the portfolio vector ξ. In the special case of $\mathbb{R}^d_+$-valued random vectors, a functional law of large numbers and a functional central limit theorem for $\widehat{\gamma}_\xi$ have been established in [27]. The present paper extends these results to the general case of random vectors in $\mathbb{R}^d$. Furthermore, the regularity assumptions are relaxed. The statistical results are shown to hold without any restrictions on the tail index α and the spectral measure Ψ. Moreover, an explicit sufficient criterion is provided for the occurrence of a nontrivial bias resulting from the estimation of the asymptotic angular distribution. Finally, aiming at Value-at-Risk in financial applications, this paper includes the estimation of the functional $\gamma_\xi^{1/\alpha}$, which characterizes the quantile asymptotics. The consistency and asymptotic normality results for $\widehat{\gamma}_\xi$ are extended to $\widehat{\gamma}_\xi^{1/\widehat{\alpha}}$.

Beyond extreme value theory, the proofs heavily rely on the theory of empirical processes with functional index (see [37]). The plug-in approach to the estimation of the functionals γ_ξ and $\gamma_\xi^{1/\alpha}$ can be applied to other functionals of Ψ and α (see Section 5). The techniques in the proofs of the functional laws of large numbers and the functional central limit theorems may be of wider interest.

The paper is structured as follows. Section 2 gives an introduction to multivariate regular variation and the characterization of portfolio loss asymptotics through the functional γ_ξ. The estimation approach and the main statistical results are presented in Section 3. The regularity assumptions underlying the statistical results are discussed in Section 4. The conclusions are drawn in Section 5, which also contains sketches of possible generalizations. Section 6 contains the proofs of the results stated in this paper.

2 MULTIVARIATE REGULAR VARIATION AND ASYMPTOTIC DEPENDENCE STRUCTURES

Consider a random vector $X = (X^{(1)}, \dots, X^{(d)})$ in $\mathbb{R}^d$ representing losses and gains generated by some assets. Focusing on the risky side, let positive component values $X^{(i)}$ represent losses, and let the gains be indicated by negative $X^{(i)}$. Then the portfolio loss is given by

$$\xi^\top X := \sum_{i=1}^{d} \xi^{(i)} X^{(i)},$$

where $\xi = (\xi^{(1)}, \dots, \xi^{(d)})$ is a vector of portfolio weights.

As a special case, this notation includes relative losses of assets $Z^{(i)}$ in a one-period model. Setting $X^{(i)} := (Z_0^{(i)} - Z_1^{(i)})/Z_0^{(i)}$, one obtains

$$\xi^\top X = \sum_{i=1}^{d} \frac{\xi^{(i)}}{Z_0^{(i)}} \bigl(Z_0^{(i)} - Z_1^{(i)}\bigr).$$

Thus $\xi^\top X$ represents the random loss generated by investing the value $\xi^{(i)}$ in the ith asset, and the relative portfolio loss equals $\xi^\top X / \sum_{i=1}^{d} \xi^{(i)}$.

Following the intuition of diversifying a unit amount of capital, let ξ be restricted to a subset H of the hyperplane

$$H_1 := \bigl\{x \in \mathbb{R}^d \colon x^{(1)} + \dots + x^{(d)} = 1\bigr\}.$$

A particularly important special case is the exclusion of negative portfolio weights (so-called short positions). The corresponding portfolio set is the unit simplex

$$\Sigma^d := \bigl\{x \in \mathbb{R}_+^d \colon x^{(1)} + \dots + x^{(d)} = 1\bigr\}.$$

Aiming at dependence of extremes, we assume throughout the paper that the probability distribution of X features a nontrivial dependence structure in the tails. To introduce the necessary notions, let us start with the definition of a regularly varying function.

A Lebesgue measurable function $f : \mathbb{R}_+ \to \mathbb{R}_+$ is called *regularly varying* (at ∞) if there exists a function $g : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\forall x \in \mathbb{R}_+, \quad \lim_{t\to\infty} \frac{f(tx)}{f(t)} = g(x).$$

It is well known that g is necessarily equal to x^β for some $\beta \in \mathbb{R}$. In the case $\beta = 0$, the function f is called *slowly varying*. Furthermore, regular variation of f is equivalent to

$$f(x) = x^\beta l(x)$$

with slowly varying l. For further details on regular variation of functions, the reader is referred to [6]. We often denote $f \in \mathcal{RV}_\beta$.

A random variable Y in $\mathbb{R}_+$ is called regularly varying with *tail index* $\alpha \geqslant 0$ if the corresponding tail probability $\bar{F}_Y(x) := 1 - F_Y(x)$ is regularly varying with index $-\alpha$:

$$\bar{F}_Y \in \mathcal{RV}_{-\alpha}.$$

For brevity and convenience, the short notation $Y \in \mathcal{RV}_{-\alpha}$ will also be used for regular variation of random variables. In the case of random variables in $\mathbb{R}$, regular variation can be considered separately for the lower and upper tails with additional balance condition (see [33, Sect. 6.5.5]).

The following definition provides the central model assumption of the paper.

DEFINITION 1. A random vector X in $\mathbb{R}^d$ is *multivariate regularly varying* if there exist a sequence $a_n \to \infty$ and a (nonzero) Radon measure ν on the Borel σ-field $\mathcal{B}([-\infty,\infty]^d \setminus \{0\})$ such that $\nu([-\infty,\infty]^d \setminus \mathbb{R}^d) = 0$ and, as $n \to \infty$,

$$n\mathbf{P}^{a_n^{-1}X} \xrightarrow{\text{v}} \nu \quad \text{on } \mathcal{B}\big([-\infty,\infty]^d \setminus \{0\}\big), \tag{2.1}$$

where $\xrightarrow{\text{v}}$ denotes *vague convergence* of Radon measures, and $\mathbf{P}^{a_n^{-1}X}$ is the probability distribution of $a_n^{-1}X$.

If X is restricted to $\mathbb{R}_+^d$, then ν is concentrated on $[0,\infty]^d \setminus \{0\}$. Therefore multivariate regular variation in this special case can also be defined by vague convergence on $\mathcal{B}([0,\infty]^d \setminus \{0\})$. For a full account of technical details related to multivariate regular variation, vague convergence, and the Borel σ-fields on the punctured spaces $[-\infty,\infty]^d \setminus \{0\}$ and $[0,\infty]^d \setminus \{0\}$, the reader is referred to [33].

It is well known that the limit measure ν obtained in (2.1) is unique except for a constant factor, has a singularity in the origin in the sense that $\nu((-\varepsilon,\varepsilon)^d) = \infty$ for any $\varepsilon > 0$, and exhibits the scaling property

$$\nu(tA) = t^{-\alpha}\nu(A) \tag{2.2}$$

for some $\alpha > 0$ and all sets $A \in \mathcal{B}([-\infty,\infty]^d \setminus \{0\})$ that are bounded away from 0.

Furthermore, (2.1) implies that $\|X\| \in \mathcal{RV}_{-\alpha}$ for any norm $\|\cdot\|$ on $\mathbb{R}^d$. The sequence a_n can always be chosen as

$$a_n := F^{\leftarrow}_{\|X\|}\left(1 - \frac{1}{n}\right),$$

where $F^{\leftarrow}_{\|X\|}$ is the quantile function of $\|X\|$. The resulting limit measure ν is normalized by

$$\nu\big(\{x \in \mathbb{R}^d \colon \|x\| > 1\}\big) = 1. \tag{2.3}$$

In addition to (2.1), we assume that the limit measure ν is nondegenerate in the following sense:

$$\nu\big(\{x \in \mathbb{R}^d \colon \big|x^{(i)}\big| > 1\}\big) > 0, \quad i = 1,\dots,d. \tag{2.4}$$

This assumption ensures that all components $X^{(i)}$ are relevant for the extremes of $\xi^\top X$. If (2.4) is satisfied in the upper tail region, i.e., if

$$\nu\big(\{x \in \mathbb{R}^d \colon x^{(i)} > 1\}\big) > 0, \quad i = 1,\dots,d,$$

then ν also characterizes the asymptotic distribution of the componentwise maxima $M_n := (M^{(1)},\dots,M^{(d)})$ obtained from an i.i.d. sequence $X_1,\dots,X_n$ via $M^{(i)} := \max\{X_1^{(i)},\dots,X_n^{(i)}\}$. In this case, one has

$$a_n^{-1}M_n \xrightarrow{\text{w}} Y \sim G$$

with the limit distribution function

$$G(y) := \exp\big(-\nu\big([-\infty,\infty]^d \setminus [-\infty,y]\big)\big), \quad y \in (0,\infty]^d.$$

Therefore ν is called *exponent measure*. For further details on the asymptotic distributions of maxima see [32] and [9].

It is also well known that the scaling property (2.2) implies a product representation of ν in polar coordinates with respect to any norm $\|\cdot\|$ on $\mathbb{R}^d$:

$$(r,s) := \tau(x) := \big(\|x\|, \|x\|^{-1}x\big).$$

The induced measure $\nu^\tau := \nu \circ \tau^{-1}$ necessarily satisfies

$$\nu^\tau = c \cdot \rho_\alpha \otimes \Psi \tag{2.5}$$

with the constant factor $c = \nu(\{x \in \mathbb{R}^d\colon \|x\| > 1\})$, the measure ρ_α on $(0,\infty]$ defined by

$$\rho_\alpha\big((x,\infty]\big) := x^{-\alpha}, \quad x \in (0,\infty],$$

and a probability measure Ψ on the unit sphere induced by $\|\cdot\|$,

$$\mathbb{S}^d_{\|\cdot\|} := \big\{s \in \mathbb{R}^d\colon \|s\| = 1\big\}.$$

The measure Ψ is called *spectral* or *angular measure* of X. In the sequel, we normalize ν according to (2.3). This entails $c = 1$ in (2.5). Using this normalization, it is easy to see that (2.1) is equivalent to

$$\mathcal{L}\big(\tau\big(t^{-1}X\big) \bigm| \|X\| > t\big) \xrightarrow{\mathrm{w}} \rho_\alpha \otimes \Psi, \quad t \to \infty, \tag{2.6}$$

on $\mathcal{B}((1,\infty] \times \mathbb{S}^d_{\|\cdot\|})$. This suggests the notion of multivariate regular variation *with tail index α and spectral measure Ψ*, abbreviated by $X \in \mathcal{MRV}_{-\alpha,\Psi}$. In the special case of $\mathbb{R}^d_+$-valued random vectors X, it may be convenient to reduce the domain of Ψ to $\mathbb{S}^d_{\|\cdot\|} \cap \mathbb{R}^d_+$.

Although the domain of Ψ depends on the norm underlying the polar coordinates, representation (2.5) is norm-independent. If (2.5) holds for some norm $\|\cdot\|$, then it is also true for any other norm $\|\cdot\|_\diamond$ that is equivalent to $\|\cdot\|$. In the following, we use the sum norm $\|x\|_1 := \sum_{i=1}^d |x^{(i)}|$ and denote by Ψ the spectral measure on the unit sphere $\mathbb{S}^d_1$ induced by $\|\cdot\|_1$.

Finally, multivariate regular variation of a random vector X is closely related to the univariate regular variation of portfolio losses $\xi^\top X$. In nondegenerate cases, $X \in \mathcal{MRV}_{-\alpha,\Psi}$ implies that $\xi^\top X \in \mathcal{RV}_{-\alpha}$ for all ξ. For further details and for the inverse, Cramér–Wold-type results we refer to [4] and [7].

The property of multivariate regular variation appears in many popular stochastic models. The examples include heavy-tailed elliptical distributions [23] and various copula models [1, 3, 17, 18]. Finally, (2.5) shows that any combination of a probability measure Ψ on $\mathbb{S}^d$ with a heavy-tailed distribution on $\mathbb{R}^d_+$ leads to a multivariate regularly varying model. Deviation from the exact product structure in the polar coordinates can be easily implemented by distortions that disappear in the tail region.

More details on regular variation of functions or random variables can be found in [4, 6, 9, 24, 32, 33].

The following lemma provides a characterization of the asymptotic portfolio losses in multivariate regularly varying models. The special case of random vectors in $\mathbb{R}^d_+$ was studied in [27]. The asymptotic portfolio risk factor γ_ξ introduced there and called *extreme risk index* of the portfolio ξ has an immediate generalization for random vectors in $\mathbb{R}^d$.

Lemma 1. *Let $X \in \mathcal{MRV}_{-\alpha,\Psi}$, $\alpha > 0$. Then*

$$\text{(a)} \quad \lim_{t\to\infty} \frac{\mathbf{P}\{\xi^\top X > t\}}{\mathbf{P}\{\|X\|_1 > t\}} = \gamma_\xi := \int_{\mathbb{S}^d_1} \big(\xi^\top s\big)_+^\alpha \,\mathrm{d}\Psi(s); \tag{2.7}$$

$$\text{(b)} \quad \lim_{u\uparrow 1} \frac{F^{\leftarrow}_{\xi^\top X}(u)}{F^{\leftarrow}_{\|X\|_1}(u)} = \gamma_\xi^{1/\alpha}. \tag{2.8}$$

Proof. See [27], Eqs. (3.1) and (2.8). □

The immediate consequence of (2.7) and (2.8) is that the functional γ_ξ characterizes the asymptotics of portfolio loss probabilities and the corresponding high loss quantiles. The limit relation (2.8) allows for an asymptotic comparison of the *Value-at-Risk* associated with different portfolio vectors ξ. The Value-at-Risk $\mathrm{VaR}_{1-\lambda}(Y)$ of a random loss Y at the level $1-\lambda$ is defined as the $1-\lambda$ quantile of Y (cf. [30]):

$$\mathrm{VaR}_{1-\lambda}(Y) := F_Y^{\leftarrow}(1-\lambda).$$

Further extensions to the asymptotic ordering of the *Expected Shortfall* $\mathrm{ES}_{1-\lambda}$ and other *spectral risk measures* are also possible [27].

3 ESTIMATION

According to (2.7), the functional γ_ξ is obtained by indexing the measure Ψ by a function $f_{\xi,\alpha}(s) := (\xi^\top x)_+^\alpha$:

$$\gamma_\xi = \Psi f_{\xi,\alpha} := \int f_{\xi,\alpha}(s)\, \mathrm{d}\Psi(s).$$

Combining an estimator $\widehat{\Psi}$ with an estimator $\widehat{\alpha}$, we obtain a plug-in estimator for γ_ξ:

$$\widehat{\gamma}_\xi := \widehat{\Psi} f_{\xi,\widehat{\alpha}}. \tag{3.1}$$

A natural estimator for the functional $\gamma_\xi^{1/\alpha}$ obtained in (2.8) would be $\widehat{\gamma}_\xi^{1/\widehat{\alpha}}$.

In the following, we consider the estimation of γ_ξ and $\gamma_\xi^{1/\alpha}$ from an i.i.d. sample $X_1, \ldots, X_n \sim X$. Working with polar coordinates of X in 1-norm, we will denote them by (R, S):

$$R := \|X\|_1, \qquad S := \|X\|_1^{-1} X.$$

Accordingly, R_i and S_i will denote the radial and angular parts of X_i for $i = 1, \ldots, n$. All estimators are based upon the subsample related to k upper order statistics $R_{n:1}, \ldots, R_{n:k}$ of the radial parts. The number $k = k(n)$ satisfies

$$k = k(n) \to \infty, \quad \frac{k}{n} \to 0 \quad \text{as } n \to \infty.$$

To avoid technicalities, we assume that the distribution F_R of the radial parts is continuous:

$$F_R \in \mathcal{C}(\mathbb{R}_+).$$

Then the sample indices $i(n,1) < \cdots < i(n,k)$ of $R_{n:1}, \ldots, R_{n:k}$ are well defined almost surely, and the corresponding angular parts can be written as $S_{i(n,1)}, \ldots, S_{i(n,k)}$. The resulting empirical estimator of Ψ is given by

$$\widehat{\Psi} = \mathbb{P}_n := \frac{1}{k} \sum_{j=1}^{k} \delta_{S_{i(n,j)}}. \tag{3.2}$$

The estimator of the tail index α is supposed to be a function of $R_{n:1}, \ldots, R_{n:k}$. Various estimation approaches are possible (see, among others, [11, 21, 31, 36]). Instead of specifying the tail index estimator $\widehat{\alpha}$, we will only impose assumptions on $\widehat{\alpha}$, such as strong consistency or asymptotic normality. This allows us to choose $\widehat{\alpha}$ according to the application.

The result cited below gives insight into the distribution structure of the extreme subsample $X_{i(n,1)}, \ldots, X_{i(n,k)}$ and the corresponding angular parts $S_{i(n,1)}, \ldots, S_{i(n,k)}$. The proof is given in [27].

Lemma 2. *Let $X \in \mathcal{MRV}_{-\alpha,\Psi}$, assume that $F_R \in \mathcal{C}(\mathbb{R}_+)$, and denote*

$$U_n := F_R(R_{n:k+1}).$$

Then, for any $u \in (0,1)$,

$$\mathcal{L}(X_{i(n,1)}, \dots, X_{i(n,k)} \mid U_n = u) = \bigotimes_{i=1}^{k} \mathcal{L}\big(X \bigm| F_R(R) > u\big). \tag{3.3}$$

An immediate consequence of (3.3) is

$$\mathcal{L}(S_{i(n,1)}, \dots, S_{i(n,k)} \mid U_n = u) = \bigotimes_{i=1}^{k} \Psi_u, \tag{3.4}$$

where

$$\Psi_u := \mathcal{L}\big(S \bigm| F_R(R) > u\big).$$

The conditional i.i.d. structure obtained in (3.4) can also be written as

$$\mathbf{P}\big\{(S_{i(n,1)}, \dots, S_{i(n,k)}) \in A\big\} = \int\limits_{[0,1]} \Psi_u^k(A)\, \mathrm{d}\mathbf{P}^{U_n}(u).$$

Here A is a Borel subset of $(\mathbb{S}_1^d)^k$, $\mathbf{P}^{U_n}$ is the probability distribution of U_n, and $\Psi_u^k := \bigotimes_{i=1}^k \Psi_u$ for $u \in (0,1)$. Moreover, since $F_R^{\leftarrow}(u) \to \infty$ for $u \uparrow 1$, multivariate regular variation of X implies

$$\Psi_u \xrightarrow{\mathrm{w}} \Psi, \quad u \uparrow 1. \tag{3.5}$$

The central results of the present paper are the uniform strong consistency and the uniform asymptotic normality of $\widehat{\gamma}_\xi$ and $\widehat{\gamma}_\xi^{1/\widehat{\alpha}}$. These properties are related to the theory of empirical measures and empirical processes. Interested readers are referred to [37].

We start with strong consistency.

Theorem 1. *Suppose that $H \subset H_1$ is compact, $k(n) \geqslant \delta n^q$ for some $q \in (0,1)$ and $\delta > 0$, and $\widehat{\alpha}$ is strongly consistent*:

$$\widehat{\alpha} \xrightarrow{\text{a.s.}} \alpha.$$

Then $\widehat{\gamma}_\xi$ is strongly consistent uniformly in $\xi \in H$:

$$\sup_{\xi \in H} |\widehat{\gamma}_\xi - \gamma_\xi| \xrightarrow{\text{a.s.}} 0. \tag{3.6}$$

Remark 1. According to [29], the n^q growth rate for $k(n)$ is a natural assumption for the strong consistency of the Hill estimator $\widehat{\alpha}_{\mathrm{H}}$. Since $\widehat{\alpha}_{\mathrm{H}}$ is the most prototypical estimator for α, this assumption does not restrict the applicability of Theorem 1.

Given (3.6), the uniform strong consistency of $\widehat{\gamma}_\xi^{1/\widehat{\alpha}}$ follows from $\widehat{\alpha} \overset{\text{a.s.}}{\to} \alpha \in (0,\infty)$ and the uniform continuity of the mapping $(t,\alpha) \mapsto t^{1/\alpha}$ on $[0,K] \times [\varepsilon, 1/\varepsilon]$ for any $K, \varepsilon \in (0,\infty)$.

Corollary 1. *The assumptions of Theorem* 1 *also imply that*

$$\sup_{\xi \in H} \bigl| \widehat{\gamma}_\xi^{1/\widehat{\alpha}} - \gamma_\xi^{1/\alpha} \bigr| \xrightarrow{\text{a.s.}} 0.$$

For asymptotic normality we need some regularity assumptions and additional notation.

Condition 1. (a) $\widehat{\alpha}$ *is asymptotically normal*:

$$\sqrt{k}(\widehat{\alpha} - \alpha) \xrightarrow{\text{w}} Y \sim \mathcal{N}\bigl(\mu_\alpha, \sigma_\alpha^2\bigr). \tag{3.7}$$

(b) *The random variable* $Y_n := \sqrt{k}(\widehat{\alpha} - \alpha)$ *and the mapping* $\mathbb{G}_n : \Omega \to l^\infty(\mathcal{F}_{H,\alpha})$ *defined by*

$$\mathbb{G}_n := \sqrt{k}(\mathbb{P}_n - \Psi_{U_n}) \tag{3.8}$$

with the random centering

$$\Psi_{U_n}(\omega) := \Psi_{U_n(\omega)}$$

are asymptotically independent. That is, for any bounded continuous functions $h_1 \in \mathcal{C}_b(\mathbb{R})$ *and* $h_2 \in \mathcal{C}_b(\mathcal{C}(\mathcal{F}_{H,\alpha}))$, *we have*

$$\lim_{n\to\infty} \mathbf{E}\bigl[h_1(Y_n) h_2(\mathbb{G}_n)\bigr] - \mathbf{E}h_1(Y_n)\mathbf{E}h_2(\mathbb{G}_n) = 0.$$

(c) *There exists a mapping* $b \in l^\infty(H)$ *such that*

$$\sup_{\xi \in H} \bigl| \sqrt{k}(\Psi_{U_n} - \Psi) f_{\xi,\alpha} - b(\xi) \bigr| \xrightarrow{\mathbf{P}} 0. \tag{3.9}$$

Remark 2. (a) Many popular estimators of the tail index α are asymptotically normal under appropriate second-order conditions specifying the convergence rate of the distribution $\mathcal{L}(t^{-1}R \mid R > t)$ for $t \to \infty$. A comprehensive elaboration on this topic can be found in [9]. For original results see, among others, [8, 11, 12, 13, 36].

(b) Condition 1(b) allows us to leave $\widehat{\alpha}$ in Theorem 2 unspecified. However, since asymptotic independence of radial and angular parts is an essential feature of multivariate regularly varying models, this assumption meets the natural intuition toward any sensible estimator $\widehat{\alpha} = \widehat{\alpha}(R_{i(n,1)}, \ldots, R_{i(n,k)})$ and the empirical process $\mathbb{G}_n$, constructed from the angular parts $S_{i(n,1)}, \ldots, S_{i(n,k)}$. In particular, the *Hill estimator* $\widehat{\alpha}_{\mathrm{H}}$, representing one of the most fundamental approaches to the estimation of the tail index α, satisfies Condition 1(b) automatically. See Lemma 4 and Corollary 2 for further details.

(c) Condition 1(c) can be understood as a second-order condition for the angular parts $S_{i(n,1)}, \ldots, S_{i(n,k)}$. Since multivariate regular variation leaves convergence rates completely unspecified, similar conditions are necessary for establishing asymptotic normality in regularly varying models. An explicit sufficient criterion for (3.9) is obtained in Lemma 3 and illustrated in Example 1.

In the following, let $\partial_{(\cdot)}$ denote the partial derivative, e.g.,

$$\partial_\alpha f_{\xi,\alpha} := \frac{\partial}{\partial \alpha} f_{\xi,\alpha}.$$

Further, let $\mathbb{G}_\Psi$ denote the *Ψ-Brownian bridge* on a function class $\mathcal{F}$, which is a tight stochastic process with index $f \in \mathcal{F}$ and multivariate Gaussian finite-dimensional marginal distributions

$$(\mathbb{G}_\Psi f_1, \ldots, \mathbb{G}_\Psi f_m) \sim \mathcal{N}(0, C). \tag{3.10}$$

The covariance structure is determined by Ψ as follows:

$$C_{i,j} = \Psi\big[(f_i - \Psi f_i)(f_j - \Psi f_j)\big] = \Psi f_i f_j - \Psi f_i \Psi f_j. \tag{3.11}$$

The tightness of $\mathbb{G}_\Psi$ implies that this process has a version with σ_Ψ-continuous paths (see [37, Sect. 2.1.2]). The *variance seminorm* σ_Ψ is defined by

$$\sigma_\Psi(f) := \big(\Psi\big[(f - \Psi f)^2\big]\big)^{1/2}. \tag{3.12}$$

Now we can state the uniform asymptotic normality of $\widehat{\gamma}_\xi$.

Theorem 2. (a) *Let $X \in \mathcal{MRV}_{-\alpha,\Psi}$, $\alpha > 0$, and assume that Condition* 1 *is satisfied. Then $\widehat{\gamma}_\xi$ is asymptotically normal uniformly in $\xi \subset H$ for compact $H \subset H_1$*:

$$\sqrt{k}(\widehat{\gamma}_\xi - \gamma_\xi) \xrightarrow{\mathrm{w}} b(\xi) + \mathbb{G}_\Psi f_{\xi,\alpha} + \Psi[\partial_\alpha f_{\xi,\alpha}]Y \quad \textit{in } l^\infty(H). \tag{3.13}$$

Here, $b(\xi)$ is the asymptotic bias term from (3.9)*, $\mathbb{G}_\Psi$ is a Ψ-Brownian bridge on $\mathcal{F}_{H,\alpha}$, and Y is the Gaussian limit in* (3.7)*. Thus Y is independent of $\mathbb{G}_\Psi$.*

(b) *Suppose that the assumptions of part* (a) *are satisfied except for Condition* 1(c)*, which is satisfied only pointwise*:

$$\sqrt{k}(\Psi_{U_n} f_{\xi_i,\alpha} - \Psi f_{\xi_i,\alpha}) \xrightarrow{\mathbf{P}} b(\xi_i) \in \mathbb{R} \tag{3.14}$$

for $\xi_1, \dots, \xi_p \in H$. Then

$$\sqrt{k}\big((\widehat{\gamma}_{\xi_1}, \dots, \widehat{\gamma}_{\xi_p}) - (\gamma_{\xi_1}, \dots, \gamma_{\xi_p})\big) \xrightarrow{\mathrm{w}} \mathcal{N}(M, C), \tag{3.15}$$

where the mean vector $M = M(\alpha, \xi_1, \dots, \xi_p)$ and the covariance matrix $C = C(\alpha, \xi_1, \dots, \xi_p)$ are given by

$$M^{(i)} = b(\xi_i) + \mu_\alpha \Psi[\partial_\alpha f_{\xi_i,\alpha}], \tag{3.16}$$

$$C_{i,j} = \Psi[f_{\xi_i,\alpha} f_{\xi_j,\alpha}] - \Psi f_{\xi_i,\alpha} \Psi f_{\xi_j,\alpha} + \sigma_\alpha^2 \Psi[\partial_\alpha f_{\xi_i,\alpha}]\Psi[\partial_\alpha f_{\xi_j,\alpha}] \tag{3.17}$$

with i, j ranging in $\{1, \dots, p\}$ and μ_α, σ_α^2 from (3.7)*.*

The asymptotic normality of $\widehat{\gamma}_\xi$ extends to $\widehat{\gamma}_\xi^{1/\widehat{\alpha}}$.

Theorem 3. *Suppose that H is compact and $\gamma_\xi \neq 0$ for all $\xi \in H$. Then*

(a) *The assumptions of Theorem* 2(a) *imply that*

$$\sqrt{k}\big(\widehat{\gamma}_\xi^{1/\widehat{\alpha}} - \gamma_\xi^{1/\alpha}\big) \xrightarrow{\mathrm{w}} c_1 Y + c_2 Z \quad \textit{in } l^\infty(H),$$

where Y is the Gaussian limit in (3.7)*, and Z is the right-hand side of* (3.13)*. The constant factors c_i are given by*

$$c_1 := -\frac{1}{\alpha^2}\gamma_\xi^{1/\alpha}\log\gamma_\xi \quad \textit{and} \quad c_2 := \frac{1}{\alpha}\gamma_\xi^{1/\alpha-1}.$$

(b) *The assumptions of Theorem* 2(b) *imply that*

$$\sqrt{k}\big(\big(\widehat{\gamma}_{\xi_1}^{1/\widehat{\alpha}}, \dots, \widehat{\gamma}_{\xi_p}^{1/\widehat{\alpha}}\big) - \big(\gamma_{\xi_1}^{1/\alpha}, \dots, \gamma_{\xi_p}^{1/\alpha}\big)\big) \xrightarrow{\mathrm{w}} \mathcal{N}\big(c_1\mu_\alpha + c_2 M, c_1^2\sigma_\alpha^2 + c_2^2 C\big)$$

with M and C defined in (3.16) *and* (3.17)*, respectively.*

4 EXAMPLES AND COMMENTS

This section is dedicated to the regularity assumptions underlying Theorem 2. We start with an explicit criterion that implies Condition 1(c).

Lemma 3. *Let $X \in \mathcal{MRV}_{-\alpha,\Psi}$, $H \subset H_1$ compact, and $F_R \in \mathcal{C}(\mathbb{R}_+)$. Suppose that*

$$\sqrt{k}(\Psi_{1-k/n} - \Psi) f_{\xi,\alpha} \to b(\xi) \quad \text{in } l^\infty(H) \tag{4.1}$$

and that the mapping $u \mapsto \Psi'_u f_{\xi,\alpha}$ with $\Psi'_u := \mathcal{L}(S \mid F_R(R) = u)$ is continuous in $u \in (0,1]$ for any $\xi \in H$. Then

$$\sqrt{k}(\Psi_{U_n} - \Psi) f_{\xi,\alpha} \xrightarrow{\mathbf{P}} b(\xi) \quad \text{in } l^\infty(H).$$

The following example illustrates Lemma 3 and shows that the angular bias term $b = b(\xi)$ depends on the choice of the extreme subsample size $k = k(n)$.

Example 1. Consider a multivariate regularly varying distribution with conditional angular distribution $\Psi'_u := \mathcal{L}(S \mid F_R(R) = u)$ given by

$$\Psi'_u := u\Psi'_1 + (1-u)\Psi'_0,$$

where Ψ'_1 and Ψ'_0 are arbitrary probability measures on $\mathcal{B}(\mathbb{S}_1^d)$. Given the continuity of the radial distribution F_R, the conditional angular distribution $\Psi_u := \mathcal{L}(S \mid F_R(R) > u)$ is equal to

$$\begin{aligned} \Psi_u &= \frac{1}{1-u}\int\limits_{(u,1)} \Psi'_v \, \mathrm{d}v = \Psi'_1 \frac{1}{1-u}\int\limits_{(u,1)} v \, \mathrm{d}v + \Psi'_0 \frac{1}{1-u}\int\limits_{(u,1)} (1-v) \, \mathrm{d}v \\ &= \frac{1+u}{2}\Psi'_1 + \frac{1-u}{2}\Psi'_0. \end{aligned}$$

This yields that the spectral measure Ψ is equal to Ψ'_1, and therefore

$$\Psi_{1-k/n} - \Psi = \frac{-k/n}{2}\Psi'_1 + \frac{k/n}{2}\Psi'_0 = \frac{k}{2n}\big(\Psi'_0 - \Psi'_1\big).$$

Hence condition (4.1) is equivalent to

$$\frac{k^{3/2}}{2n}\big(\Psi'_0 - \Psi'_1\big) f_{\xi,\alpha} \to b(\xi) \quad \text{in } l^\infty(H).$$

Consequently, (4.1) is satisfied if $k^{3/2}/(2n) \to \lambda \in [0,\infty)$. The asymptotic bias term $b(\xi)$ appearing in Theorem 2 is given by

$$b(\xi) = \lambda\big(\Psi'_0 - \Psi'_1\big) f_{\xi,\alpha}.$$

In particular, $b(\xi)$ is nonzero for $\lambda > 0$.

Another point that is worth a discussion is the asymptotic independence of the normalized estimation error $Y_n = \sqrt{k}(\widehat{\alpha} - \alpha)$ and the empirical process $\mathbb{G}_n$ stated in Condition 1(b). As already highlighted in Remark 2, this condition is rather natural in the framework of multivariate regular variation and is automatically satisfied by the Hill estimator. The rest of the current section provides a proof for this assertion.

The *Hill estimator* [21], defined as

$$\widehat{\alpha}_{\mathrm{H}} := \left(\frac{1}{k}\sum_{i=1}^{k} \log\left(\frac{R_{n:i}}{R_{n:k+1}}\right)\right)^{-1},$$

is one of the earliest and most popular estimators for the tail index α of a heavy-tailed distribution. Denoting $\widetilde{R}_{i(n,j)} := R_{i(n,j)}/R_{n:k+1}$, one obtains the representation

$$\widehat{\alpha}_{\mathrm{H}}^{-1} = \frac{1}{k}\sum_{j=1}^{k} \log \widetilde{R}_{i(n,j)}.$$

Hence the tuple $(\widehat{\alpha}_{\mathrm{H}}^{-1}, \mathbb{P}_n f_{\xi,\alpha})$ can be written as

$$\bigl(\widehat{\alpha}_{\mathrm{H}}^{-1}, \mathbb{P}_n f_{\xi,\alpha}\bigr) = (\widetilde{\mathbb{P}}_n \tilde{l}, \widetilde{\mathbb{P}}_n \tilde{f}_{\xi,\alpha}) \tag{4.2}$$

with the empirical measure $\widetilde{\mathbb{P}}_n$ defined by

$$\widetilde{\mathbb{P}}_n := \frac{1}{k}\sum_{i=1}^{k} \delta_{(\widetilde{R}_{i(n,j)}, S_{i(n,j)})}$$

and the functional indices $\tilde{l}$, $\tilde{f}_{\xi,\alpha}$ defined by

$$\tilde{l}(r,s) := \log(r) \quad \text{and} \quad \tilde{f}_{\xi,\alpha}(r,s) := f_{\xi,\alpha}(s).$$

Recall that $R_{n:k+1} = F_R^{\leftarrow}(U_n)$ $\mathbf{P}$-a.s. for continuous F_R. Consequently, Lemma 2 yields

$$\mathcal{L}\bigl((\widetilde{R}_{i(n,1)}, S_{i(n,1)}), \ldots, (\widetilde{R}_{i(n,k)}, S_{i(n,k)}) \bigm| U_n = u\bigr) = \bigotimes_{i=1}^{k} \widetilde{\mathbf{P}}_u,$$

where

$$\widetilde{\mathbf{P}}_u := \mathcal{L}\biggl(\frac{R}{F_R^{\leftarrow}(u)}, S \biggm| F_R(R) > u\biggr). \tag{4.3}$$

Representation (4.2) shows that the asymptotic independence of the normalized estimation error $Y_n := \sqrt{k}(\widehat{\alpha}_{\mathrm{H}} - \alpha)$ and the empirical process $\mathbb{G}_n$ assumed in Condition 1(b) is related to the asymptotic behavior of the empirical process

$$\widetilde{\mathbb{G}}_n := \sqrt{k}(\widetilde{\mathbb{P}}_n - \widetilde{\mathbf{P}}_{U_n}) \tag{4.4}$$

with random centering $\widetilde{\mathbf{P}}_{U_n}(\omega) := \widetilde{\mathbf{P}}_{U_n(\omega)}$ and functional index

$$f \in \widetilde{\mathcal{F}}_{H,\alpha} := \{\tilde{l}\} \cup \{\tilde{f}_{\xi,\alpha}\colon \xi \in H\}.$$

The following lemma states the weak convergence of the empirical process $\widetilde{\mathbb{G}}_n$ to a Gaussian process.

Lemma 4. *Let $X \in \mathcal{MRV}_{-\alpha,\Psi}$, $\alpha > 0$, $H \subset H_1$ compact, and $F_R \in \mathcal{C}(\mathbb{R}_+)$. Then the empirical process $\widetilde{\mathbb{G}}_n$ defined in* (4.4) *satisfies*

$$\widetilde{\mathbb{G}}_n \xrightarrow{\mathrm{w}} \mathbb{G}_{\rho_\alpha \otimes \Psi} \quad \textit{in } \mathcal{C}(\widetilde{\mathcal{F}}_{H,\alpha}). \tag{4.5}$$

The final result of this section is obtained by combination of Lemma 4 with the Delta method.

Corollary 2. *Suppose that the conditions of Lemma* 4 *are satisfied and*

$$\sqrt{k}\big(\widetilde{\mathrm{P}}_{U_n}\tilde{l} - \alpha^{-1}\big) \xrightarrow{\mathbf{P}} b \in \mathbb{R}. \tag{4.6}$$

Then $\widehat{\alpha}_{\mathrm{H}}$ *is asymptotically normal and satisfies Condition* 1(b)*, i.e., the random variable* $Y_n := \sqrt{k}(\widehat{\alpha}_{\mathrm{H}} - \alpha)$ *is asymptotically independent from* $\mathbb{G}_n$*.*

Remark 3. It is well known that condition (4.6) can be ensured by strengthening the regular variation of the radial part R by a second-order condition and adding a regularity condition on the sequence $k = k(n)$. For technical details on the asymptotic normality of tail index estimates, we refer to [9].

5 CONCLUSIONS AND GENERALIZATIONS

The approach to the comparison of extremal portfolio losses proposed in [27] for multivariate regularly varying vectors in $\mathbb{R}^d_+$ has been extended to $\mathbb{R}^d$. The statistical results from [27] have also been extended to this more general case. Analogous results have been obtained for the estimator $\widehat{\gamma}_\xi^{1/\widehat{\alpha}}$.

Moreover, the conditions underlying the asymptotic normality results have been significantly relaxed by dropping all restrictions on α and Ψ. A thorough discussion of these conditions has been provided, including a sufficient criterion for the appearance of the nontrivial angular bias $b(\xi)$. If the model has exact product structure, then the angular bias is zero. In particular, this is the case for centered multivariate t-distributions. Models featuring asymptotically vanishing distortion of the product structure may have a nontrivial angular bias, depending on the decay rate of the distortion.

The combination of extreme value theory with the theory of empirical processes presented here is also applicable to other problems in the area of extremal dependence. For instance, instead of the portfolio excess sets

$$\big\{x \in \mathbb{R}^d_+ \colon \xi^\top x > t\big\},$$

one may be interested in the asymptotic probabilities of the sets

$$\Big\{x \in \mathbb{R}^d_+ \colon \bigvee_i \xi^{(i)} x^{(i)} > t\Big\}$$

with direction parameter $\xi \in H \subset \Sigma^d$. These sets indicate that at least one of the weighted components $\xi^{(i)} X^{(i)}$ exceeds t. Analogously to Lemma 1, one obtains that the asymptotic probabilities of these sets for $t \to \infty$ can be quantified by the functional $\Psi g_{\xi,\alpha}$ with $g_{\xi,\alpha}(s) = (\bigvee_i \xi^{(i)} s^{(i)})^\alpha$. Similarly to (2.8), the asymptotic dependence factor for the corresponding quantiles is equal to $(\Psi g_{\xi,\alpha})^{1/\alpha}$.

If asymptotic independence is excluded, the same arguments can be applied to the asymptotic probabilities of the directed joint excess events $(x \in \mathbb{R}^d_+ \colon \bigwedge_i \xi^{(i)} x^{(i)} > t)$. This leads to similar limit functionals with the integrand $g_{\xi,\alpha}(s) = (\bigwedge_i \xi^{(i)} s^{(i)})^\alpha$.

Combining a semiparametric estimation approach and the techniques of the proof presented in Section 6, it is straightforward to derive functional laws of large numbers and functional central limit theorems similar to Theorems 1–3 and Corollary 1.

6 PROOFS

6.1 Empirical processes with functional index

The estimator $\widehat{\gamma}_\xi$ proposed in (3.1) is obtained by indexing the empirical measure $\mathbb{P}_n$ defined in (3.2) with a random element $f_{\xi,\widehat{\alpha}}$ of the function class

$$\mathcal{F}_H := \big\{f_{\xi,\alpha} \colon \xi \in H,\ \alpha \in (0,\infty)\big\},$$

where $H \subset H_1$ is a compact set of admissible portfolio vectors. Later on, we will see that the consistency of the estimator $\widehat{\alpha}$ and smoothness of the parameterization $\alpha \mapsto f_{\xi,\alpha}$ allow us to reduce the index set of the empirical measure $\mathbb{P}_n$ and the empirical process $\mathbb{G}_n$ defined in (3.8) to the function class

$$\mathcal{F}_{H,\alpha} := \{f_{\xi,\alpha}\colon \xi \in H\}$$

with $\alpha \in (0,\infty)$ being the true tail index.

The asymptotic normality of $\widehat{\gamma}_\xi$ can be viewed as a special version of the Donsker theorem. Let $\mathbb{P}_{k,\Psi}$ denote the empirical measure corresponding to k i.i.d. random variables with probability distribution Ψ:

$$\mathbb{P}_{k,\Psi} := \frac{1}{k}\sum_{i=1}^{k}\delta_{Y_i}, \quad Y_1,\dots,Y_k \text{ i.i.d.} \sim \Psi. \tag{6.1}$$

The corresponding *empirical process* $\mathbb{G}_{k,\Psi}$ is defined as

$$\mathbb{G}_{k,\Psi} := \sqrt{k}(\mathbb{P}_{k,\Psi} - \Psi).$$

A class $\mathcal{F}$ of measurable functions is called *pre-Gaussian* if there exists a tight Ψ-Brownian bridge $\mathbb{G}_\Psi$ on $\mathcal{F}$. A class $\mathcal{F}$ is called *Donsker* if the Donsker theorem holds for $\mathbb{G}_{k,\Psi}$ uniformly in $f \in \mathcal{F}$:

$$\mathbb{G}_{k,\Psi} \xrightarrow{\text{w}} \mathbb{G}_\Psi \quad \text{in } l^\infty(\mathcal{F}),\ k \to \infty. \tag{6.2}$$

The pre-Gaussian and Donsker properties of $\mathcal{F}$ guarantee the existence of a probability space $(\Omega', \mathcal{A}', \mathbf{P}')$ and a tight, Borel-measurable mapping $\mathbb{G}_\Psi : \Omega' \to l^\infty(\mathcal{F})$ satisfying (3.10), (3.11), and (6.2).

The notion of weak convergence in $l^\infty(\mathcal{F})$ is understood according to [37]. Based on *outer expectations* and *outer probabilities*, this extended notion allows one to consider $\mathbb{P}_n$ and $\mathbb{G}_n$ as mappings from the probability space Ω into $l^\infty(\mathcal{F})$, although the measurability in l^∞ is not available in general (see [5, Sect. 18]). In the special case where $\mathbb{P}_n$ and $\mathbb{G}_n$ are measurable, the extended notions of stochastic convergence (weak, in probability, or almost sure) coincide with the standard ones. As it will be shown below, the problem considered here is of this kind. This allows us to apply Donsker theorems from [37] and obtain the weak convergence of empirical measures in the classical sense.

However, standard Donsker theorems for i.i.d. samples cannot be applied to the subsample $S_{i(n,1)}, \dots, S_{i(n,k)}$ directly since the random variables $S_{i(n,1)}, \dots, S_{i(n,k)}$ are not necessarily independent (although they are conditionally independent given $U_n = u$; see Lemma 2). Moreover, the probability distribution of each $S_{i(n,j)}$ varies with n. Thus uniform convergence results for and $\mathbb{G}_n f_{\xi,\widehat{\alpha}}$ demand a special version of the Donsker theorem that takes into account the structure of the underlying probability distribution. This result is stated in Lemma 8, after a series of auxiliary results.

We start with an outline of some useful facts.

Remark 4. (a) The mapping $(\xi, s, \alpha) \mapsto f_{\xi,\alpha}(s)$ is continuous, and hence uniformly continuous on a compact domain. This implies that any function class $\mathcal{F}_{H,I} := \{f_{\xi,\alpha}\colon \xi \in H,\ \alpha \in I\}$ with compact H and compact $I \subset (0,\infty)$ is uniformly bounded. Moreover, such $\mathcal{F}_{H,I}$ are compact in $(\mathcal{C}(\mathbb{S}_1^d), \|\cdot\|_\infty)$. In particular, this is the case for all function classes $\mathcal{F}_{H,\alpha} := \{f_{\xi,\alpha}\colon \xi \in H\}$ with compact H. The same arguments apply to the partial derivatives $\partial_\alpha f_{\xi,\alpha}$ and $\partial_\alpha^2 f_{\xi,\alpha}$ and the corresponding function classes $\partial_\alpha \mathcal{F}_{H,I}$ and $\partial_\alpha^2 \mathcal{F}_{H,I}$.

(b) It is obvious that any probability measure Ψ on $\mathbb{S}_1^d$ satisfies $|\Psi f - \Psi g| \leqslant \|f - g\|_\infty$ for $f, g \in \mathcal{C}(\mathbb{S}_1^d)$. Thus the mapping $f \mapsto \Psi f$ is Lipschitz with factor 1.

Recall the conditional angular distribution $\Psi_u = \mathcal{L}(S \mid F_R(R) > u)$ defined for $u \in [0,1)$ (see Lemma 2). Motivated by (3.5), we introduce the extended notation $\Psi_1 := \Psi$. The subsequent lemma provides the continuity of the parameterization $u \mapsto \Psi_u$, which is essential to the measurability of the random centering Ψ_{U_n} in (3.8).

Lemma 5. *Let $X \in \mathcal{MRV}_{-\alpha,\Psi}$ with $\alpha \in (0,\infty)$, F_R continuous, and $H \subset H_1$ compact. Then*

(a) *the mappings $u \mapsto \Psi_u f_{\xi,\alpha}$ and $u \mapsto \Psi_u[\partial_\alpha f_{\xi,\alpha}]$ are continuous in $u \in [0,1]$ for any $\xi \in H$;*

(b) *the measure Ψ_u converges to Ψ in l^∞:*

$$\|\Psi_u - \Psi\|_{\mathcal{F}^*} := \sup_{f\in\mathcal{F}^*} |\Psi_u f - \Psi f| \to 0, \quad u \uparrow 1,$$

for $\mathcal{F}^* = \mathcal{F}_{H,\alpha}$ and $\mathcal{F}^* = \partial_\alpha \mathcal{F}_{H,\alpha} := \{\partial_\alpha f_{\xi,\alpha}\colon \xi \in H\}$.

Proof. Part (a). The continuity of F_R implies that $F_R(R) \sim \mathrm{unif}(0,1)$, and therefore

$$\Psi_u f = \frac{\mathbf{E}[f(S)\mathbf{1}\{F_R(R) > u\}]}{1-u}$$

for $u < 1$. Moreover, $\mathbf{P}\{F_R(R) = u\} = 0$ implies that

$$\mathbf{1}\{F_R(R) > u_n\} \xrightarrow{\text{a.s.}} \mathbf{1}\{F_R(R) > u\}$$

for any sequence $u_n \to u < 1$. Thus $\Psi_{u_n} f \to \Psi_u f$ follows from the dominated convergence theorem. The continuity of $u \mapsto \Psi_u f$ in $u = 1$ is an immediate consequence of the weak convergence $\Psi_u \xrightarrow{\text{w}} \Psi = \Psi_1$ established in (3.5).

Part (b). According to Remark 4(b), the mapping $f \mapsto \Psi f$ is Lipschitz(1) for all Ψ. Hence the family $\{\Psi_u\colon u \in [0,1]\}$ can be considered as an equicontinuous subset of $\mathcal{C}(\mathcal{F}^*)$. The uniform convergence $\Psi_{u_n} f \to \Psi_u f$ for $f \in \mathcal{F}^*$ follows from the compactness of $\mathcal{F}^*$ stated in Remark 4(b) and the pointwise convergence in part (a). □

The following result guarantees that the random measures involved in the proof of Theorem 2 can be treated as random variables in $\mathcal{C}(\mathcal{F}^*)$.

Lemma 6. *The empirical measures $\mathbb{P}_n$ and $\mathbb{P}_{k,\Psi}$, the random measures Ψ_{U_n}, and the empirical processes $\mathbb{G}_n$ and $\mathbb{G}_{k,\Psi}$ are Borel-measurable mappings in $\mathcal{C}(\mathcal{F}^*)$ for $\mathcal{F}^* = \mathcal{F}_{H,\alpha}$ and $\mathcal{F}^* = \partial_\alpha \mathcal{F}_{H,\alpha}$.*

Proof. According to Remark 4(a), $\mathcal{F}^*$ is a compact subset of $\mathcal{C}(\mathbb{S}_1^d)$. Moreover, Remark 4(b) implies that the mappings $f \mapsto \mathbb{P}_n(\omega)f$, $f \mapsto \mathbb{P}_{k,\Psi}(\omega)f$, $f \mapsto \Psi_{U_n}(\omega)f$, $f \mapsto \mathbb{G}_n(\omega)f$, and $f \mapsto \mathbb{G}_{k,\Psi}(\omega)f$ are continuous in $f \in \mathcal{F}^*$ for any $\omega \in \Omega$. Since $\mathcal{F}^*$ is compact, a mapping $\omega \mapsto \phi(\omega)$ from Ω into $\mathcal{C}(\mathcal{F}^*)$ is Borel-measurable if $\omega \mapsto (\phi(\omega))(f)$ is measurable for all $f \in \mathcal{F}^*$ [37, Ex. 1.5.1]. Thus it suffices to show the measurability of the random variables $\mathbb{P}_n f$, $\mathbb{P}_{k,\Psi} f$, and $\Psi_{U_n} f$ for every $f \in \mathcal{F}^*$. The measurability of $\mathbb{G}_n f$, and $\mathbb{G}_{k,\Psi} f$ is a trivial consequence.

It is easy to see that the mappings $\omega \mapsto \mathbb{P}_n(\omega)f$ and $\omega \mapsto \mathbb{P}_{k,\Psi}(\omega)f$ for $f \in \mathcal{F}^*$ are measurable by construction (see (3.2) and (6.1)). The measurability of $\omega \mapsto \Psi_{U_n}(\omega)f = \Psi_{U_n(\omega)}f$ follows immediately from the measurability of U_n and the continuity of the mapping $u \mapsto \Psi_u f$ established in Lemma 5(a). □

A function class $\mathcal{F}$ is called *universally* Donsker or pre-Gaussian if the corresponding property holds uniformly for all probability measures on the sample space.

Lemma 7. *Let $H \subset H_1$ be compact, and $\alpha \in (0,\infty)$. Then the function class $\mathcal{F}_{H,\alpha}$ is universally Donsker and pre-Gaussian.*

Proof. It is obvious that all functions $f \in \mathcal{F}_{H,\alpha}$ are measurable and uniformly bounded (see Remark 4(a)). Hence the constant function

$$F(s) := \mathbf{1}_{\mathbb{S}_1^d}(s) \sup_{f\in\mathcal{F}_{H,\alpha}} \|f\|_\infty$$

can serve as an *envelope function* for $\mathcal{F}_{H,\alpha}$. That is, we have $F(s) \geqslant |f(s)|$ for all $s \in \mathbb{S}_1^d$ and $f \in \mathcal{F}_{H,\alpha}$.

The separability of $\mathcal{F}_{H,\alpha}$ in $l^\infty(\mathbb{S}_1^d)$ implies that $\mathcal{F}_{H,\alpha}$ is *universally Ψ-measurable*, i.e., Ψ-measurable for any probability measure Ψ on $\mathcal{B}(\mathbb{S}_1^d)$ [37, Def. 2.3.3]. Another consequence of the separability of $\mathcal{F}_{H,\alpha}$ is the separability of the function classes $\{f-g\colon f,g\in\mathcal{F}_{H,\alpha},\ \|f-g\|_{\Psi,2}<\delta\}$ and $\{(f-g)^2\colon f,g\in\mathcal{F}_{H,\alpha}\}$. Therefore these function classes are universally Ψ-measurable, and [37, Thm. 2.8.3] yields that $\mathcal{F}_{H,\alpha}$ is universally Donsker and pre-Gaussian if the following *uniform entropy condition* is satisfied:

$$\int_0^\infty \sup_{Q\in\mathcal{Q}} \sqrt{\log N\big(\varepsilon\|F\|_{Q,2},\mathcal{F}_{H,\alpha},L^2(Q)\big)}\,\mathrm{d}\varepsilon<\infty. \tag{6.3}$$

Here $\mathcal{Q}$ denotes the set of all finitely discrete probability measures, and $N(\varepsilon,\mathcal{F}_{H,\alpha},L^2(Q))$ is the number of ε-balls in $L^2(Q)$ needed to cover $\mathcal{F}_{H,\alpha}$. An ε-ball around f is defined as $\{g\in L^2(Q)\colon \|g-f\|_{Q,2}<\varepsilon\}$.

Notice that $\mathcal{F}_{H,\alpha}$ is covered by a singe ball of size $\|F\|_{Q,2}$:

$$\forall f\in\mathcal{F}_{H,\alpha},\quad \|f\|_{Q,2}\leqslant\|F\|_{Q,2}.$$

This allows us to reduce (6.3) to

$$\int_0^1 \sup_{Q\in\mathcal{Q}} \sqrt{\log N\big(\varepsilon\|F\|_{Q,2},\mathcal{F}_{H,\alpha},L^2(Q)\big)}\,\mathrm{d}\varepsilon<\infty. \tag{6.4}$$

It is obvious that $|(\xi_1^\top s)_+-(\xi_2^\top s)_+|\leqslant\|\xi_1-\xi_2\|_\infty$ for $s\in\mathbb{S}_1^d$. This implies that the mapping $\phi_\alpha:\xi\mapsto f_{\xi,\alpha}$ from H to $\mathcal{C}(\mathbb{S}_1^d)$ is Lipschitz for $\alpha\geqslant 1$, whilst for $\alpha\in(0,1)$, we obtain that

$$\big\|\phi_\alpha(\xi_1)-\phi_\alpha(\xi_2)\big\|_\infty=\sup_{s\in\mathbb{S}_1^d}\big|(\xi_1^\top s)_+^\alpha-(\xi_2^\top s)_+^\alpha\big|\leqslant\sup_{s\in\mathbb{S}_1^d}\big|(\xi_1^\top s)_+-(\xi_2^\top s)_+\big|^\alpha\leqslant\|\xi_1-\xi_2\|^\alpha.$$

Hence, if we cover H by m balls of radius δ, the set $\mathcal{F}_{H,\alpha}=\phi_\alpha(H)$ is covered by m balls of radius $c\delta^{\alpha\wedge 1}$ (in $\|\cdot\|_\infty$) for some $c>0$. Being a compact subset of $\mathbb{R}^d$, H can be covered by $m=O(\delta^{-d})$ balls of radius δ. Therefore $\mathcal{F}_{H,\alpha}$ can be covered by $O(\varepsilon^{-d/(\alpha\wedge1)})$ balls of size ε. As any ball in $\|\cdot\|_\infty$ is smaller than the ball in any $L^2(Q)$ metric with the same center and radius, the polynomial bound $O(\varepsilon^{-d/(\alpha\wedge1)})$ also holds for the covering number N in (6.4) uniformly in Q. Hence the integrability condition (6.4) is satisfied. □

Now we can prove the weak convergence of the empirical process $\mathbb{G}_n$ defined in (3.8).

Lemma 8. *Suppose that $X\in\mathcal{MRV}_{-\alpha,\Psi}$ and that $H\subset H_1$ is compact. Then the empirical process $\mathbb{G}_n=\sqrt{k}(\mathbb{P}_n-\Psi_{U_n})$ satisfies*

$$\mathbb{G}_n\xrightarrow{\mathrm{w}}\mathbb{G}_\Psi\quad \textit{in}\ \mathcal{C}(\mathcal{F}_{H,\alpha}).$$

Proof. According to Lemma 6, $\mathbb{G}_n$ is a Borel-measurable mapping into $\mathcal{C}(\mathcal{F}_{H,\alpha})$. Thus weak convergence is understood in the classical way, and it suffices to show that

$$\lim_{n\to\infty}\mathbf{E}h(\mathbb{G}_n)=\mathbf{E}h(\mathbb{G}_\Psi) \tag{6.5}$$

for any function $h\in\mathcal{C}_b(\mathcal{C}(\mathcal{F}_{H,\alpha}))$. Applying Lemma 2, we obtain that

$$\mathbf{E}h(\mathbb{G}_n)=\mathbf{E}\big[\mathbf{E}\big[h(\mathbb{G}_n)\;\big|\;U_n\big]\big]=\mathbf{E}\big[\bar{h}_n(U_n)\big]$$

with $\bar{h}_n(u) := \mathbf{E}h(\mathbb{G}_{k,\Psi_u})$ for $k = k(n)$. Thus we have to show that

$$\mathbf{E}\bar{h}_n(U_n) \to \bar{h}_\infty(1) := \mathbf{E}h(\mathbb{G}_\Psi). \tag{6.6}$$

As h is bounded and $U_n \overset{\text{a.s.}}{\to} 1$, the continuous mapping principle [5, Thm. 5.5] would yield (6.6) if we can prove that $\bar{h}_n(u_n) \to \bar{h}_\infty(1)$ for any sequence $u_n \uparrow 1$. Thus, in order to verify (6.5), it suffices to show that

$$\mathbb{G}_{k,\Psi_k} \xrightarrow{\text{w}} \mathbb{G}_\Psi \quad \text{in } \mathcal{C}(\mathcal{F}_{H,\alpha}) \tag{6.7}$$

for $\Psi_k := \Psi_{u_k}$ and $u_k \uparrow 1$.

Recall that $\Psi_u \overset{\text{w}}{\to} \Psi$ as $u \uparrow 1$ according to (3.5). Hence we have $\Psi_k \overset{\text{w}}{\to} \Psi$. Furthermore, Lemma 7 states that the function class $\mathcal{F}_{H,\alpha}$ is universally Donsker and pre-Gaussian, and $\mathcal{F}_{H,\alpha}$ is uniformly bounded according to Remark 4(a). Thus, according to [37, Lemma 2.8.7], condition (6.7) is satisfied if

$$\sup_{f_1,f_2 \in \mathcal{F}_{H,\alpha}} \left|\sigma_{\Psi_k}(f_1 - f_2) - \sigma_\Psi(f_1 - f_2)\right| \to 0, \tag{6.8}$$

where σ_Ψ is the variance seminorm introduced in (3.12). Denote $\mathcal{G} := \mathcal{F}_{H,\alpha} - \mathcal{F}_{H,\alpha}$. As $g \mapsto \sigma_\Psi(g)$ is continuous on $\mathcal{G}$ for any Ψ, condition (6.8) can also be written as $\sigma_{\Psi_k} \to \sigma_\Psi$ in $\mathcal{C}(\mathcal{G})$. Since $\mathcal{G}$ is compact (as a continuous image of the compact $\mathcal{F}_{H,\alpha} \times \mathcal{F}_{H,\alpha}$), $\sigma_{\Psi_k} \to \sigma_\Psi$ in $\mathcal{C}(\mathcal{G})$ is equivalent to $\sigma_{\Psi_k} g_k \to \sigma_\Psi g$ for $g_k \to g$. This, however, easily follows from $\Psi_k g_k^2 \to \Psi g^2$ and $\Psi_k g_k \to \Psi g$ due to $\Psi_k \overset{\text{w}}{\to} \Psi$. $\square$

6.2 Proofs of the main results

This subsection contains the proofs of Theorems 1, 2, and 3.

Proof of Theorem 1. Since $\widehat{\alpha} \overset{\text{a.s.}}{\to} \alpha$, we only need to consider $\omega \in \Omega$ for which $\widehat{\alpha}$ is consistent. In this case, we have that

$$\sup_{\xi \in H} \|f_{\xi,\widehat{\alpha}} - f_{\xi,\alpha}\|_\infty \to 0$$

as $n \to \infty$ because the mapping $(\xi, \alpha, s) \mapsto f_{\xi,\alpha}(s)$ is continuous and H and $\mathbb{S}_1^d$ are compact. Furthermore, each $\widehat{\Psi}$ is Lipschitz(1) as a mapping from $\mathcal{C}(\mathbb{S}_1^d)$ to $\mathbb{R}$ (see Remark 4(b)). This yields

$$|\widehat{\gamma}_\xi - \gamma_\xi| \leqslant \|f_{\xi,\widehat{\alpha}} - f_{\xi,\alpha}\|_\infty + |\widehat{\Psi} f_{\xi,\alpha} - \Psi f_{\xi,\alpha}|.$$

Hence it suffices to prove that $\widehat{\Psi} f_{\xi,\alpha} \to \Psi f_{\xi,\alpha}$ uniformly in $\xi \in H$. However, since the function class $\mathcal{F}_{H,\alpha} := \{f_{\xi,\alpha}\colon \xi \in H\}$ is compact in $(\mathcal{C}(\mathbb{S}_1^d), \|\cdot\|_\infty)$ and all $\widehat{\Psi}$ are Lipschitz(1), the uniform convergence of $\widehat{\Psi}$ on $\mathcal{F}_{H,\alpha}$ is equivalent to the pointwise one.

It is obvious that

$$|\widehat{\Psi} f_{\xi,\alpha} - \Psi f_{\xi,\alpha}| \leqslant |\widehat{\Psi} f_{\xi,\alpha} - \Psi_{U_n} f_{\xi,\alpha}| + |\Psi_{U_n} f_{\xi,\alpha} - \Psi f_{\xi,\alpha}|.$$

Thus, due to $U_n \overset{\text{a.s.}}{\to} 1$ and $\Psi_{u_n} \overset{\text{w}}{\to} \Psi$, it suffices to show that

$$|\widehat{\Psi} f_{\xi,\alpha} - \Psi_{U_n} f_{\xi,\alpha}| \xrightarrow{\text{a.s.}} 0. \tag{6.9}$$

According to Lemma 2,

$$\mathcal{L}(\widehat{\Psi} f_{\xi,\alpha} - \Psi_{U_n} f_{\xi,\alpha} \mid U_n = u) = \mathcal{L}\left(\sum_{j=1}^{k} \frac{1}{k}\big(f_{\xi,\alpha}(Y_j) - \Psi_u f_{\xi,\alpha}\big)\right)$$

with i.i.d. random vectors $Y_1, \dots, Y_k \sim \Psi_u$. Denote $Z_j := (f_{\xi,\alpha}(Y_j) - \Psi_u f_{\xi,\alpha})$. Since $f_{\xi,\alpha}$ is bounded, we have $|Z_j| \leqslant M$ for some $M > 0$. Thus Hoeffding's inequality (see [22, Thm. 2]) yields

$$\mathbf{P}\biggl\{\biggl|\sum_{j=1}^{k} \frac{1}{k} Z_j\biggr| > \varepsilon\biggr\} \leqslant \exp\biggl(\frac{-k\varepsilon^2}{2M^2}\biggr)$$

universally for all Ψ_u, and hence

$$\sum_{n=1}^{\infty} \mathbf{P}\bigl\{|\widehat{\Psi} f_{\xi,\alpha} - \Psi_{U_n}| > \varepsilon\bigr\} \leqslant \sum_{n=1}^{\infty} \exp\biggl(\frac{-k(n)\varepsilon^2}{2M^2}\biggr). \tag{6.10}$$

By assumption we have that $k(n) \geqslant \delta n^q$ for some $q \in (0,1)$ and $\delta > 0$. This gives a finite sum in (6.10), and the Borel–Cantelli lemma yields (6.9). □

Proof of Theorem 2. Part (a). Consider the decomposition

$$\begin{aligned}\sqrt{k}(\widehat{\gamma}_\xi - \gamma_\xi) &= \sqrt{k}(\mathbb{P}_n f_{\xi,\widehat{\alpha}} - \Psi f_{\xi,\alpha}) = \sqrt{k}\bigl(\mathbb{P}_n[f_{\xi,\widehat{\alpha}} - f_{\xi,\alpha}] + (\mathbb{P}_n - \Psi_{U_n}) f_{\xi,\alpha} + (\Psi_{U_n} - \Psi) f_{\xi,\alpha}\bigr) \\ &= \mathbb{P}_n\bigl[\sqrt{k}(f_{\xi,\widehat{\alpha}} - f_{\xi,\alpha})\bigr] + \mathbb{G}_n f_{\xi,\alpha} + \sqrt{k}(\Psi_{U_n} - \Psi) f_{\xi,\alpha}. \end{aligned} \tag{6.11}$$

First it should be noted that Condition 1(c) postulates

$$\sqrt{k}(\Psi_{U_n} - \Psi) f_{\xi,\alpha} \xrightarrow{\mathbf{P}} b(\xi) \quad \text{in } l^\infty(H)$$

and Lemma 8 implies

$$\mathbb{G}_n f_{\xi,\alpha} \xrightarrow{\mathrm{w}} \mathbb{G}_\Psi f_{\xi,\alpha} \quad \text{in } \mathcal{C}(H).$$

Thus it suffices to consider the first term in (6.11). Recall that the function class $\partial_\alpha^2 \mathcal{F}_{H,\alpha}$ is uniformly bounded according to Remark 4(a). Hence, for any sequence $y_n \to y$ in $\mathbb{R}$,

$$\sqrt{k}\bigl(f_{\xi,\alpha + y_n\sqrt{k}}(s) - f_{\xi,\alpha}(s)\bigr) - \partial_\alpha f_{\xi,\alpha}(s) y_n \to 0$$

uniformly in ξ and s, and thus in the space $\mathcal{C}(H \times \mathbb{S}_1^d)$. This implies that

$$\widehat{g}_{\xi,k} := \sqrt{k}(f_{\xi,\widehat{\alpha}} - f_{\xi,\alpha}) - \sqrt{k}\partial_\alpha f_{\xi,\alpha}(\widehat{\alpha} - \alpha) \xrightarrow{\mathbf{P}} 0$$

in $\mathcal{C}(\mathbb{S}_1^d)$ uniformly in $\xi \in H$, and $|\mathbb{P}_n \widehat{g}| \leqslant \mathbb{P}_n \|g\|_\infty$ yields

$$\mathbb{P}_n\bigl[\sqrt{k}(f_{\xi,\widehat{\alpha}} - f_{\xi,\alpha})\bigr] - \sqrt{k}(\widehat{\alpha} - \alpha)\mathbb{P}_n \partial_\alpha f_{\xi,\alpha} \xrightarrow{\mathbf{P}} 0$$

in $\mathcal{C}(H)$. Thus it suffices to show that

$$\sqrt{k}(\widehat{\alpha} - \alpha)\mathbb{P}_n \partial_\alpha f_{\xi,\alpha} \xrightarrow{\mathrm{w}} Y \Psi \partial_\alpha f_{\xi,\alpha} \tag{6.12}$$

in $\mathcal{C}(H)$ with a Gaussian random variable $Y \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$. Since $\mathbb{P}_n \partial_\alpha f_{\xi,\alpha} \xrightarrow{\mathbf{P}} \Psi \partial_\alpha f_{\xi,\alpha}$ uniformly in $\xi \in H$ (cf. Lipschitz(1) property of $\mathbb{P}_n$ and compactness of $\partial_\alpha \mathcal{F}_{H,\alpha}$ stated in Remark 4), (6.12) follows from the asymptotic normality of $\widehat{\alpha}$. Finally, the asymptotic independence of the random variable $Y_n := \sqrt{k}(\widehat{\alpha} - \alpha)$ and the empirical process $\mathbb{G}_n$ stated in Condition 1(b) yields the result (3.13).

Part (b). This result is merely the finite-dimensional version of part (a). It is easy to see that replacing assumption (3.9) by (3.14) affects only the last term in (6.11) and results in an exchange of the uniform convergence to $b(\xi)$ by a pointwise version. Hence the pointwise asymptotic normality (3.15) of $\widehat{\gamma}_\xi$ follows immediately along the lines of the proof of part (a). □

Proof of Theorem 3. Part (a). Condition 1(b) and Theorem 2(a) imply the joint convergence

$$\sqrt{k}(\widehat{\gamma}_\xi - \gamma_\xi, \widehat{\alpha} - \alpha) \xrightarrow{\mathrm{w}} (Z, Y),$$

where Y is the Gaussian limit in (3.7), independent of Z, and Z is the random mapping in $l^\infty(H)$ that appears on the right-hand side of (3.13):

$$Z = Z(\xi) := b(\xi) + \mathbb{G}_n f_{\xi,\alpha} + \Psi[\partial_\alpha f_{\xi,\alpha}]. \tag{6.13}$$

Moreover, the continuity of γ_ξ, the compactness of H, and the assumption $\gamma_\xi \neq 0$ imply that $\gamma_\xi \in (\varepsilon, 1/\varepsilon)$ for all $\xi \in H$ with some $\varepsilon > 0$. The constant ε can always be chosen such that $\alpha \in (\varepsilon, 1/\varepsilon)$. Since the mapping $\phi : (\gamma, \alpha) \mapsto \gamma^{1/\alpha}$ is $\mathcal{C}^2$ on $(0, \infty)^2$, the first-order Taylor approximation of $\phi(\widehat{\gamma}_\xi, \widehat{\alpha}) - \phi(\gamma_\xi, \alpha)$ is uniform for $(\widehat{\gamma}_\xi, \widehat{\alpha}) \in [\varepsilon, 1/\varepsilon]^2$, and (6.13) implies that

$$\sqrt{k}\big(\phi(\widehat{\gamma}_\xi, \widehat{\alpha}) - \phi(\gamma_\xi, \alpha)\big) \xrightarrow{\mathrm{w}} \partial_\alpha \phi(\widehat{\gamma}_\xi, \alpha) Y + \partial_t \phi(\widehat{\gamma}_\xi, \alpha) Z.$$

The result follows from $\partial_\alpha \phi(t, \alpha) = -\alpha^{-2} t^{1/\alpha} \log t$ and $\partial_t \phi(t, \alpha) = \alpha^{-1} t^{1/\alpha - 1}$.

Part (b) is analogous, with obvious calculations. □

6.3 Discussion of the regularity assumptions

Proof of Lemma 3. Due to (4.1), it suffices to show that

$$(\Psi_{U_n} - \Psi_{1-k/n}) f_{\xi,\alpha} = o_{\mathbf{P}}\left(\frac{1}{\sqrt{k}}\right)$$

uniformly in $\xi \in H$. Recall the notation $\Psi'_u := \mathcal{L}(S \mid F_R(R) = u)$. Due to $F_R(R) \sim \mathrm{unif}(0, 1)$, we have that

$$\Psi_u f = \frac{1}{1-u} \int_{(u,1)} \Psi'_v f \,\mathrm{d}v$$

for $u \in [0, 1)$ and $f \in \mathcal{F}_{H,\alpha}$. Thus the continuity of the mapping $u \mapsto \Psi'_u f_{\xi,\alpha}$ entails the differentiability of $u \mapsto \Psi_u f_{\xi,\alpha}$:

$$\frac{\partial}{\partial u} \Psi_u f_{\xi,\alpha} = \frac{1}{(1-u)^2} \int_{(u,1)} \Psi'_v f_{\xi,\alpha} \,\mathrm{d}v - \frac{1}{1-u} \Psi'_u f_{\xi,\alpha} = \frac{1}{1-u} \big(\Psi_u - \Psi'_u\big) f_{\xi,\alpha}.$$

Hence the mean-value theorem yields

$$(\Psi_{U_n} - \Psi_{1-k/n}) f_{\xi,\alpha} = \frac{1}{1-u^*} \big(\Psi_{u^*} - \Psi'_{u^*}\big) f_{\xi,\alpha} \left(U_n - \left(1 - \frac{k}{n}\right)\right) \tag{6.14}$$

for some u^* between $(1-k/n)$ and U_n. It is well known [35] that the random variable $U_n = F_R(R_{n:k+1})$ satisfies

$$\frac{U_n - (1-k/n)}{\sqrt{k}/n} \xrightarrow{\mathrm{w}} \mathcal{N}(0,1).$$

This implies $U_n - (1-k/n) = O_{\mathbf{P}}(\sqrt{k}/n)$, and therefore $1 - u^* = k/n + O_{\mathbf{P}}(\sqrt{k}/n)$. Consequently, (6.14) yields

$$(\Psi_{U_n} - \Psi_{1-k/n}) f_{\xi,\alpha} = \frac{O_{\mathbf{P}}(\sqrt{k}/n)}{k/n + O_{\mathbf{P}}(\sqrt{k}/n)} \bigl(\Psi_{u^*} - \Psi'_{u^*}\bigr) f_{\xi,\alpha} = O_{\mathbf{P}}\biggl(\frac{1}{\sqrt{k}}\biggr) \bigl(\Psi_{u^*} - \Psi'_{u^*}\bigr) f_{\xi,\alpha}.$$

Hence it suffices to show that

$$\sup_{\xi \in H} \bigl| \bigl(\Psi_{u^*} - \Psi'_{u^*}\bigr) f_{\xi,\alpha} \bigr| = o_{\mathbf{P}}(1). \tag{6.15}$$

It is easy to see that

$$\begin{aligned}\sup_{\xi \in H} \bigl| \bigl(\Psi_{u^*} - \Psi'_{u^*}\bigr) f_{\xi,\alpha} \bigr| &= \sup_{\xi \in H} \biggl| \frac{1}{1-u^*} \int\limits_{(u^*,1)} \bigl(\Psi'_v - \Psi'_{u^*}\bigr) f_{\xi,\alpha} \,\mathrm{d}v \biggr| \\ &\leqslant \sup_{v \in [u^*,1],\, \xi \in H} \bigl| \bigl(\Psi'_v - \Psi'_{u^*}\bigr) f_{\xi,\alpha} \bigr|. \end{aligned} \tag{6.16}$$

Recall that the mapping $u \mapsto \Psi'_u f_{\xi,\alpha}$ is continuous on $(0,1]$ for $\xi \in H$ by assumption. Furthermore, the mapping $\xi \mapsto \Psi'_u f_{\xi,\alpha}$ is continuous on H for $u \in (0,1]$ since each Ψ'_u is a probability measure and $\xi \mapsto f_{\xi,\alpha}$ is continuous in $l^\infty(\mathbb{S}_1^d)$ (see Remark 4). Hence the mapping $(u,\xi) \mapsto \Psi'_u f_{\xi,\alpha}$ is continuous on $(0,1] \times H$ and therefore uniformly continuous on $[\varepsilon,1] \times H$ for any $\varepsilon > 0$. This implies that

$$\sup_{v \in [u,1],\, \xi \in H} \bigl| \bigl(\Psi'_v - \Psi'_u\bigr) f_{\xi,\alpha} \bigr| \to 0, \quad u \uparrow 1. \tag{6.17}$$

Finally, since u^* is always chosen between $(1-k/n)$ and U_n, we obtain $u^* \xrightarrow{\mathbf{P}} 1$. Thus (6.15) is a consequence of (6.16) and (6.17). $\square$

Proof of Lemma 4. According to (2.6), $X \in \mathcal{MRV}_{-\alpha,\Psi}$ means

$$\widetilde{\mathbf{P}}_u \xrightarrow{\mathrm{w}} \rho_\alpha \otimes \Psi, \quad u \uparrow 1,$$

where $\widetilde{\mathbf{P}}_u$ is the conditional distribution introduced in (4.3). Thus the weak convergence (4.5) can be considered as an extension of Lemma 8 and proved by adapting the proof of Lemma 8 to $\widetilde{\mathbb{G}}_n$ and $\widetilde{\mathcal{F}}_{H,\alpha}$. An envelope function for the function class $\widetilde{\mathcal{F}}_{H,\alpha}$ is given by

$$\widetilde{F}(r,s) := \max\bigl(\log(r), F(s)\bigr),$$

where F is an envelope function of the function class $\mathcal{F}_{H,\alpha}$. Since F is bounded, the integrability of $\widetilde{F}$ and $\widetilde{F}^2$ depends only on the integrability of $\tilde{l}$ and $\tilde{l}^2$. Let π_1 denote the projection on the first component: $\pi_1(r,s) := r$. Then we have

$$\widetilde{\mathbf{P}}_u \tilde{l} = \int\limits_{(1,\infty)} \log(r) \,\mathrm{d}\widetilde{\mathbf{P}}_u^{\pi_1}(r) = \int\limits_{(1,\infty)} \int\limits_{(1,\infty)} \mathbf{1}\{v < r\}\, v^{-1} \,\mathrm{d}v \,\mathrm{d}\widetilde{\mathbf{P}}_u^{\pi_1}(r)$$

$$= \int\limits_{(1,\infty)} v^{-1} \int\limits_{(1,\infty)} \mathbf{1}\{v < r\}\, \mathrm{d}\widetilde{\mathbf{P}}_u^{\pi_1}(r)\, \mathrm{d}v = \int\limits_{(1,\infty)} v^{-1} \mathbf{P}\Big\{ \frac{R}{F_R^{\leftarrow}(y)} > v \,\Big|\, R > F_R^{\leftarrow}(u) \Big\}\, \mathrm{d}v$$
$$= \int\limits_{(1,\infty)} v^{-1} \frac{\mathbf{P}\{R > vt\}}{\mathbf{P}\{R > t\}}\, \mathrm{d}v \tag{6.18}$$

with $t = t(u) := F_R^{\leftarrow}(u)$. Thus the regular variation of the random variable R and Karamata's theorem [6, Thm. 1.5.11(ii)] yield $\widetilde{\mathbf{P}}_u \tilde{l} < \infty$ for $u \in (0,1)$ and

$$\lim_{u \uparrow 1} \widetilde{\mathbf{P}}_u \tilde{l} = \lim_{t \to \infty} \int\limits_{(1,\infty)} v^{-1} \frac{\mathbf{P}\{R > vt\}}{\mathbf{P}\{R > t\}}\, \mathrm{d}v = \int\limits_{(1,\infty)} v^{-1-\alpha}\, \mathrm{d}v = \frac{1}{\alpha}. \tag{6.19}$$

Analogously to (6.18), we obtain

$$\widetilde{\mathbf{P}}_u \tilde{l}^2 = \int\limits_{(1,\infty)} 2v^{-1} \log(v) \frac{\mathbf{P}\{R > vt\}}{\mathbf{P}\{R > t\}}\, \mathrm{d}v.$$

Now the generalized Karamata theorem [9, Thm. B.1.12.2] yields $\widetilde{\mathbf{P}}_u \tilde{l}^2 < \infty$ for $u \in (0, u)$ and

$$\lim_{u \uparrow 1} \widetilde{\mathbf{P}}_u \tilde{l}^2 = \lim_{t \to \infty} \int\limits_{(1,\infty)} 2v^{-1} \log(v) \frac{\mathbf{P}\{R > vt\}}{\mathbf{P}\{R > t\}}\, \mathrm{d}v = \int\limits_{(1,\infty)} 2v^{-1-\alpha} \log(v)\, \mathrm{d}v = \frac{2}{\alpha^2}. \tag{6.20}$$

Hence $\widetilde{F}^2$ is integrable. This allows us to extend the uniform Donsker and pre-Gaussian properties established for the function class $\mathcal{F}_{H,\alpha}$ in Lemma 7 to the function class $\widetilde{\mathcal{F}}_{H,\alpha}$. Furthermore, (6.20) allows us to extend Lemma 8 to $\widetilde{\mathbb{G}}_n$ and $\widetilde{\mathcal{F}}_{H,\alpha}$ by verifying condition (6.8) for $\widetilde{\mathcal{F}}_{H,\alpha}$. That is, it suffices to show that

$$\sup_{f_1, f_2 \in \widetilde{\mathcal{F}}_{H,\alpha}} \big| \sigma_{\widetilde{\mathbf{P}}_k}(f_1 - f_2) - \sigma_{\rho_\alpha \otimes \Psi}(f_1 - f_2) \big| \to 0 \tag{6.21}$$

with $\widetilde{\mathbf{P}}_k := \widetilde{\mathbf{P}}_{u_k}$ for an arbitrary sequence $u_k \uparrow 1$.

Denote $g := f_1 - f_2$. Since $|t_1^{1/2} - t_2^{1/2}| \leqslant |t_1 - t_2|^{1/2}$ for $t_1, t_2 \geqslant 0$, a sufficient criterion for (6.21) is

$$\sup_{g \in \widetilde{\mathcal{F}}_{H,\alpha} - \widetilde{\mathcal{F}}_{H,\alpha}} \big| \sigma^2_{\widetilde{\mathbf{P}}_k}(g) - \sigma^2(g) \big| \to 0.$$

As the case $g = \tilde{f}_{\xi_1,\alpha} - \tilde{f}_{\xi_2,\alpha}$ is already covered by (6.8), we only have to consider $g = \tilde{l} - \tilde{f}_{\xi,\alpha}$. It is easy to see that

$$\begin{aligned} \big| \sigma^2_{\widetilde{\mathbf{P}}_k}(g) - \sigma^2_{\rho_\alpha \otimes \Psi}(g) \big| &= \big| \widetilde{\mathbf{P}}_k g^2 - (\widetilde{\mathbf{P}}_k g)^2 - (\rho_\alpha \otimes \Psi) g^2 + \big((\rho_\alpha \otimes \Psi) g \big)^2 \big| \\ &\leqslant \big| (\widetilde{\mathbf{P}}_k - \rho_\alpha \otimes \Psi) g^2 \big| + \big| (\widetilde{\mathbf{P}}_k - \rho_\alpha \otimes \Psi) g \big| \big| (\widetilde{\mathbf{P}}_k + \rho_\alpha \otimes \Psi) g \big|. \end{aligned}$$

Note also that $|g| \leqslant |\tilde{l}| + |\tilde{f}_{\xi,\alpha}|$, $g^2 \leqslant 2\tilde{l}^2 + 2\tilde{f}^2_{\xi,\alpha}$, and $f^2_{\xi,\alpha} \in \mathcal{F}_{H,\alpha^2}$. Hence, due to Lemma 5(b), we only need to verify that

$$\big| (\widetilde{\mathbf{P}}_k - \rho_\alpha \otimes \Psi) \tilde{l}^2 \big| \to 0 \quad \text{and} \quad \big| (\widetilde{\mathbf{P}}_k - \rho_\alpha \otimes \Psi) \tilde{l} \big| \to 0 \quad \text{as } k \to \infty$$

for arbitrary α. This follows directly from (6.19) and (6.20). □

Proof of Corollary 2. Denote $\vartheta := \alpha^{-1}$ and $\widehat{\vartheta}_{\mathrm{H}} := \widehat{\alpha}_{\mathrm{H}}^{-1}$. Then the Taylor expansion yields

$$Y_n = \sqrt{k}\big(\widehat{\vartheta}_{\mathrm{H}}^{-1} - \vartheta^{-1}\big) = -\vartheta^{-2}\sqrt{k}(\widehat{\vartheta}_{\mathrm{H}} - \vartheta) + \vartheta^{*-3}\sqrt{k}(\widehat{\vartheta}_{\mathrm{H}} - \vartheta)^2$$

for some ϑ^* between $\widehat{\vartheta}_{\mathrm{H}}$ and ϑ. Hence, due to $\widehat{\vartheta}_{\mathrm{H}} = \widetilde{\mathbb{P}}_n\tilde{l}$, we obtain

$$Y_n = -\vartheta^{-2}\sqrt{k}(\widetilde{\mathbb{P}}_n\tilde{l} - \vartheta)\tilde{l} + \vartheta^{*-3}\sqrt{k}(\widetilde{\mathbb{P}}_n\tilde{l} - \vartheta)^2. \tag{6.22}$$

Furthermore, assumption (4.6) and Lemma 4 imply that

$$\sqrt{k}(\widetilde{\mathbb{P}}_n\tilde{l} - \vartheta) = \widetilde{\mathbb{G}}_n\tilde{l} + \sqrt{k}(\widetilde{\mathbf{P}}_{U_n}\tilde{l} - \vartheta) \xrightarrow{\mathrm{w}} \mathcal{N}\big(b, \sigma_\vartheta^2\big) \tag{6.23}$$

with

$$\sigma_\vartheta^2 := (\rho_\alpha \otimes \Psi)\tilde{l}^2 - \big((\rho_\alpha \otimes \Psi)\tilde{l}\big)^2 = \alpha^{-2}.$$

This yields $\vartheta^* \xrightarrow{\mathbf{P}} \vartheta$ and

$$\vartheta^{*-3}\sqrt{k}(\widetilde{\mathbb{P}}_n\tilde{l} - \vartheta)^2 = o_{\mathbf{P}}(1). \tag{6.24}$$

Applying (6.23) and (6.24) to (6.22), we obtain the asymptotic normality of $\widehat{\alpha}$:

$$Y_n = \sqrt{k}(\widehat{\alpha}_{\mathrm{H}} - \alpha) \xrightarrow{\mathrm{w}} Y \sim \mathcal{N}\big(-b\alpha^2, \alpha^2\big).$$

Now consider the asymptotic independence of Y_n and $\mathbb{G}_n$. Since $Y_n \xrightarrow{\mathrm{w}} Y$ and $\mathbb{G}_n \xrightarrow{\mathrm{w}} \mathbb{G}_\Psi$, the asymptotic independence of Y_n and $\mathbb{G}_n$ is equivalent to the joint convergence

$$(Y_n, \mathbb{G}_n) \xrightarrow{\mathrm{w}} (Y, \mathbb{G}_\Psi) \tag{6.25}$$

with independent Y and $\mathbb{G}_\Psi$. Recall $\mathbb{G}_n f_{\xi,\alpha} = \widetilde{\mathbb{G}}_n\tilde{f}_{\xi,\alpha}$, (6.22), (6.23), and Lemma 4. This implies the joint convergence (6.25). The independence of the limits follows from the covariance structure of $\mathbb{G}_{\rho_\alpha\otimes\Psi}$. □

Acknowledgments. I thank RiskLab, ETH Zürich, for financial support. I would also like to thank Paul Embrechts and Bikramjit Das for several discussions related to the paper. Last but not least, I would like to express my gratitude to the anonymous referee, whose conscientious reading and insightful comments helped to improve this article.

REFERENCES

1. S. Alink, M. Löwe, and M.V. Wüthrich, Diversification of aggregate dependent risks, *Insur. Math. Econ.*, **35**(1):77–95, 2004.
2. A. Araujo and E. Giné, *The Central Limit Theorem for Real and Banach Valued Random Variables*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York, Chichester, Brisbane, 1980.
3. P. Barbe, A.-L. Fougères, and C. Genest, On the tail behavior of sums of dependent risks, *Astin Bull.*, **36**(2):361–373, 2006.
4. B. Basrak, R.A. Davis, and T. Mikosch, A characterization of multivariate regular variation, *Ann. Appl. Probab.*, **12**(3):908–920, 2002.
5. P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons Inc., New York, 1968.

6. N.H. Bingham, C.M. Goldie, and J.L. Teugels, *Regular Variation*, Encycl. Math. Appl., Vol. 27, Cambridge Univ. Press, Cambridge, 1987.
7. J. Boman and F. Lindskog, Support theorems for the Radon transform and Cramér–Wold theorems, *J. Theor. Probab.*, **22**(3):683–710, 2009.
8. R. Davis and S.I. Resnick, Tail estimates motivated by extreme value theory, *Ann. Stat.*, **12**(4):1467–1487, 1984.
9. L. de Haan and A. Ferreira, *Extreme Value Theory*, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2006.
10. L. de Haan and S.I. Resnick, Limit theory for multivariate sample extremes, *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, **40**(4):317–337, 1977.
11. A.L.M. Dekkers, J.H.J. Einmahl, and L. de Haan, A moment estimator for the index of an extreme-value distribution, *Ann. Stat.*, **17**(4):1833–1855, 1989.
12. H. Drees, Refined Pickands estimators of the extreme value index, *Ann. Stat.*, **23**(6):2059–2080, 1995.
13. H. Drees, A. Ferreira, and L. de Haan, On maximum likelihood estimation of the extreme value index, *Ann. Appl. Probab.*, **14**(3):1179–1201, 2004.
14. J.H.J. Einmahl, L. de Haan, and X. Huang, Estimating a multidimensional extreme-value distribution, *J. Multivariate Anal.*, **47**(1):35–47, 1993.
15. J.H.J. Einmahl, L. de Haan, and V.I. Piterbarg, Nonparametric estimation of the spectral measure of an extreme value distribution, *Ann. Stat.*, **29**(5):1401–1423, 2001.
16. J.H.J. Einmahl and J. Segers, Maximum empirical likelihood estimation of the spectral measure of an extreme-value distribution, *Ann. Stat.*, **37**(5B):2953–2989, 2009.
17. P. Embrechts, D.D. Lambrigger, and M.V. Wüthrich, Multivariate extremes and the aggregation of dependent risks: Examples and counter-examples, *Extremes*, **12**(2):107–127, 2009.
18. P. Embrechts, J. Nešlehová, and M.V. Wüthrich, Additivity properties for value-at-risk under Archimedean dependence and heavy-tailedness, *Insur. Math. Econ.*, **44**(2):164–169, 2009.
19. M. Falk, J. Hüsler, and R.-D. Reiss, *Laws of Small Numbers: Extremes and Rare Events*, DMV Semin., Vol. 23, Birkhäuser, Basel, 1994.
20. G. Gudendorf and J. Segers, Extreme-value copulas, in P. Bickel, P. Diggle, S. Fienberg, U. Gather, I. Olkin, S. Zeger, P. Jaworski, F. Durante, W.K. Härdle, and T. Rychlik (Eds.), *Copula Theory and Its Applications*, Lect. Notes Stat., Vol. 198, Springer, Berlin, Heidelberg, 2010, pp. 127–145.
21. B.M. Hill, A simple general approach to inference about the tail of a distribution, *Ann. Stat.*, **3**(5):1163–1174, 1975.
22. W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Am. Stat. Assoc.*, **58**(301):13–30, 1963.
23. H. Hult and F. Lindskog, Multivariate extremes, aggregation and dependence in elliptical distributions, *Adv. Appl. Probab.*, **34**(3):587–608, 2002.
24. H. Hult and F. Lindskog, Regular variation for measures on metric spaces, *Publ. Inst. Math., Nouv. Sér.*, **80(94)**:121–140, 2006.
25. C. Klüppelberg, G. Kuhn, and L. Peng, Estimating the tail dependence function of an elliptical distribution, *Bernoulli*, **13**(1):229–251, 2007.
26. G. Mainik and P. Embrechts, Diversification in heavy-tailed portfolios: Properties and pitfalls, preprint, 2012, available from: `http://www.math.ethz.ch/~gmainik/Papers/mrv-diversification.pdf` [cited 12 June 2012].

27. G. Mainik and L. Rüschendorf, On optimal portfolio diversification with respect to extreme risks, *Finance Stoch.*, **14**:593–623, 2010.

28. Y. Malevergne and D. Sornette, *Extreme Financial Risks*, Springer, Berlin, 2006.

29. D.M. Mason, Laws of large numbers for sums of extreme values, *Ann. Probab.*, **10**(3):754–764, 1982.

30. A.J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management*, Princeton Series in Finance, Princeton Univ. Press, Princeton, NJ, 2005.

31. J. Pickands III, Statistical inference using extreme order statistics, *Ann. Stat.*, **3**:119–131, 1975.

32. S.I. Resnick, *Extreme Values, Regular Variation, and Point Processes*, Appl. Probab., Vol. 4, Springer, New York, 1987.

33. S.I. Resnick, *Heavy-Tail Phenomena*, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2007.

34. R. Schmidt and U. Stadtmüller, Non-parametric estimation of tail dependence, *Scand. J. Stat.*, **33**(2):307–335, 2006.

35. N.V. Smirnov, Limit distributions for the terms of a variational series, *Tr. Mat. Inst. Steklova*, **25**:3–60, 1949 (in Russian).

36. R.L. Smith, Estimating tails of probability distributions, *Ann. Stat.*, **15**(3):1174–1207, 1987.

37. A.W. van der Vaart and J.A. Wellner, *Weak Convergence and Empirical Processes*, Springer Series in Statistics, Springer, New York, 1996. Corrected 2nd printing 2000.